Bayesian Inference of Model Error in Imprecise Models

Nicolas Leoni, Pietro Marco Congedo, Olivier Le Maitre, Maria-Giovanna Rodio

DE LA RECHERCHE À L'INDUSTRIE

BAYESIAN INFERENCE OF MODEL ERROR IN IMPRECISE MODELS

January 2021, ECCOMAS Congress (online)

Nicolas Leoni [1 2] Pietro Congedo [2] Olivier Le Maître [3] Maria-Giovanna Rodio [1]

[1] CEA Saclay, DES/ISAS/DM2S/STMF/LMSF [2] Inria Saclay-Île-de-France [3] CNRS, CMAP, Palaiseau

Contents

Contents

Numerical simulation in nuclear safety

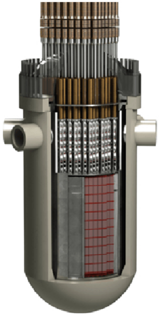


Image source : Framatome

- A reliable simulation of the reactor is needed for risk prevention ;
- Physical models exist for each component of the reactor,
- These models are calibrated separately then used together,
- Modelling assumptions can lead to inaccurate predictions.

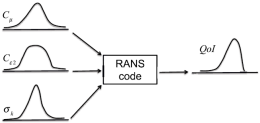
To make more realistic predictions :
Need to estimate uncertainties linked to modeling assumptions, namely ”model error”.

Contents

Calibration of a computer code

Use observations obtained from an experiment to find out the best values of coefficients in the numerical model.

It is an **inverse problem**.
Bayesian framework : the parameters are unknown and **random**.
They follow a probability distribution, from **prior** (before seeing the data) to **posterior** (after seeing the data).

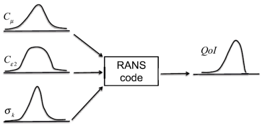


(a) Direct problem

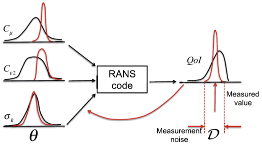
Calibration of a computer code

Use observations obtained from an experiment to find out the best values of coefficients in the numerical model.

It is an **inverse problem**.
Bayesian framework : the parameters are unknown and **random**.
They follow a probability distribution, from **prior** (before seeing the data) to **posterior** (after seeing the data).



(a) Direct problem



(b) Inverse problem

Image source : [XC19]

Uncertainty sources in numerical simulations

Traditional decomposition : ([KO01])

- **Parametric uncertainty** : due to the imperfect knowledge of coefficients in the model.
- **Model error** : due to modeling assumptions to construct the physical model.
- **Experimental uncertainty** : results from inexact measurements of quantities of interest.
- **Numerical error** : results from solving a physical model with numerical methods (discrete schemes, finite elements, . . .).

Uncertainty sources in numerical simulations

Traditional decomposition : ([KO01])

- **Parametric uncertainty** : due to the imperfect knowledge of coefficients in the model.
- **Model error** : due to modeling assumptions to construct the physical model.
- **Experimental uncertainty** : results from inexact measurements of quantities of interest.
- **Numerical error** : results from solving a physical model with numerical methods (discrete schemes, finite elements, . . .).

In Uncertainty Quantification studies, **model error** is often not included.

Uncertainty sources in numerical simulations

Traditional decomposition : ([KO01])

- **Parametric uncertainty** : due to the imperfect knowledge of coefficients in the model.
- **Model error** : due to modeling assumptions to construct the physical model.
- **Experimental uncertainty** : results from inexact measurements of quantities of interest.
- **Numerical error** : results from solving a physical model with numerical methods (discrete schemes, finite elements, . . .).

In Uncertainty Quantification studies, **model error** is often not included.

In this work : How can we accurately estimate **model error** and take it into account in the result of a computer code ?

Problem formulation

Calibration with model error [KO01] :

$$y_{obs}(x) = f(x, \theta^{best}) + z(x) + \epsilon_{exp} \tag{1}$$

- No numerical error.
- Model error : defined as difference between reality and model *taken at its best parameters*.
- What is the ”best parameters” value ?

Problem formulation

Calibration with model error [KO01] :

$$y_{obs}(x) = f(x, \theta^{best}) + z(x) + \epsilon_{exp} \tag{1}$$

- No numerical error.
- Model error : defined as difference between reality and model *taken at its best parameters*.
- What is the "best parameters" value ?

Calibration with flexible prior [Plu17] :

$$y_{obs}(x) = f(x, \theta) + z(x, \theta) + \epsilon_{exp} \tag{2}$$

- Model bias depends on model parameters.

Common hypotheses : $\epsilon_{exp} \sim N(0, \sigma^2_{exp})$; $z|\psi \sim GP(\mu, c_\psi)$; $z \perp\!\!\!\perp \epsilon_{exp}$.

With previous hypotheses, the likelihood function is gaussian :

$$\mathbf{y}_{obs}|\boldsymbol{\theta}, \psi \sim \mathrm{N}(\mathbf{f}_{\boldsymbol{\theta}}, \mathbf{C}_{\psi} + \sigma^2_{\text{exp}}\mathbf{I}_n). \tag{3}$$

And the full posterior distribution is given by Bayes' Theorem :

$$\mathrm{p}(\boldsymbol{\theta}, \psi|\mathbf{y}_{obs}) \propto \mathrm{p}(\boldsymbol{\theta}, \psi)\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta}, \psi). \tag{4}$$

Equation (4) is too expensive to compute in general.

With previous hypotheses, the likelihood function is gaussian :

$$\mathbf{y}_{obs}|\boldsymbol{\theta}, \psi \sim \mathrm{N}(\mathbf{f}_{\boldsymbol{\theta}}, \mathbf{C}_{\psi} + \sigma_{\exp}^2 \mathbf{I}_n). \tag{3}$$

And the full posterior distribution is given by Bayes' Theorem :

$$\mathrm{p}(\boldsymbol{\theta}, \psi|\mathbf{y}_{obs}) \propto \mathrm{p}(\boldsymbol{\theta}, \psi)\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta}, \psi). \tag{4}$$

Equation (4) is too expensive to compute in general.
KOH approximation :
A single value of hyperparameters is estimated with maximum a posteriori :

$$\psi_{MAP} = \arg\max_{\boldsymbol{\psi}} \; p(\psi|\mathbf{y}_{obs}). \tag{5}$$

With previous hypotheses, the likelihood function is gaussian :

$$\mathbf{y}_{obs}|\boldsymbol{\theta},\boldsymbol{\psi} \sim \mathrm{N}(\mathbf{f}_{\boldsymbol{\theta}},\mathbf{C}_{\boldsymbol{\psi}} + \sigma^2_{\mathrm{exp}}\mathbf{I}_n). \qquad (3)$$

And the full posterior distribution is given by Bayes' Theorem :

$$\mathrm{p}(\boldsymbol{\theta},\boldsymbol{\psi}|\mathbf{y}_{obs}) \propto \mathrm{p}(\boldsymbol{\theta},\boldsymbol{\psi})\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta},\boldsymbol{\psi}). \qquad (4)$$

Equation (4) is too expensive to compute in general.

KOH approximation :

A single value of hyperparameters is estimated with maximum a posteriori :

$$\boldsymbol{\psi}_{MAP} = \underset{\boldsymbol{\psi}}{\arg\max}\; p(\boldsymbol{\psi}|\mathbf{y}_{obs}). \qquad (5)$$

Full Maximum Posterior approximation :

Multiple values of hyperparameters are estimated :

$$\boldsymbol{\psi}_{FMP}(\boldsymbol{\theta}) = \underset{\boldsymbol{\psi}}{\arg\max}\; \mathrm{p}(\boldsymbol{\psi})\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta},\boldsymbol{\psi}). \qquad (6)$$

The posterior density of parameters is then estimated by :

$$\mathrm{p}(\boldsymbol{\theta}|\mathbf{y}_{obs}) \propto \mathrm{p}(\boldsymbol{\theta})\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta}, \boldsymbol{\psi} = \boldsymbol{\psi}_{FMP}(\boldsymbol{\theta})) \quad (7)$$

Calibration problem (cont.)

The posterior density of parameters is then estimated by :

$$\mathrm{p}(\boldsymbol{\theta}|\mathbf{y}_{obs}) \propto \mathrm{p}(\boldsymbol{\theta})\mathrm{p}(\mathbf{y}_{obs}|\boldsymbol{\theta}, \boldsymbol{\psi} = \boldsymbol{\psi}_{FMP}(\boldsymbol{\theta})) \tag{7}$$

The model can be used for prediction at a new point x^* :

$$\mathrm{p}(f(x^*)|\mathbf{y}_{obs}) = \int_{\boldsymbol{\theta}} \mathrm{p}(f(x^*)|\boldsymbol{\theta}, \mathbf{y}_{obs})\mathrm{p}(\boldsymbol{\theta}|\mathbf{y}_{obs})\,\mathrm{d}\boldsymbol{\theta} \tag{8}$$

Algorithm for FMP calibration

Algorithm 1: FMP calibration

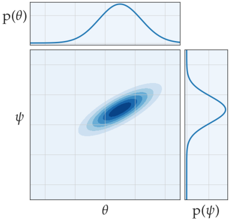
- Build a DoE $\{\boldsymbol{\theta}_i\}$ in the model parameters space;
- Run the computer model to get $\{f(\boldsymbol{\theta}_i)\}$ and perform the optimisations to get $\{\psi_{FMP}(\boldsymbol{\theta}_i)\}$ (eq. (6));
- Store the optimized values and multiply by p($\boldsymbol{\theta}$) to get the posterior density p($\boldsymbol{\theta}|\mathbf{y}_{obs}$);
- The posterior density is used to compute new predictions.

- More optimisations are required compared to traditional methods.
- Optimisations are independent and can be run in parallel.
- In this work we use a full grid sampling. Integrals are evaluated by quadrature.

Contents

Unimodal Gaussian Posterior

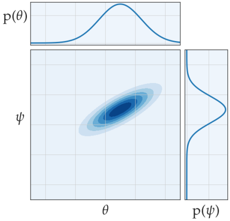
Let us assume that the joint posterior p($\theta, \psi | \mathbf{y}_{obs}$) is gaussian.



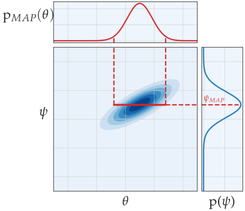
(a) Full Bayes

Unimodal Gaussian Posterior

Let us assume that the joint posterior $\mathrm{p}(\boldsymbol{\theta}, \psi | \mathbf{y}_{obs})$ is gaussian.

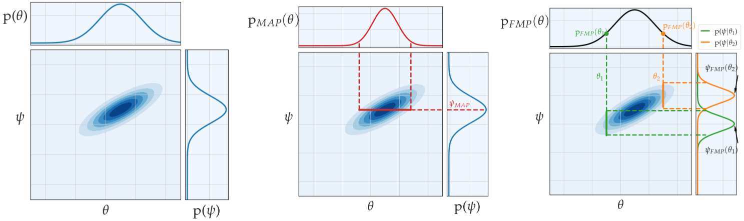


(a) Full Bayes



(b) MAP approximation

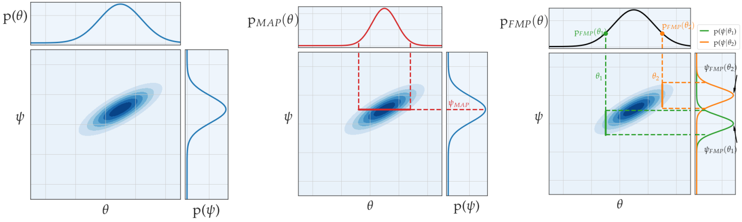
Let us assume that the joint posterior $p(\boldsymbol{\theta}, \psi | \mathbf{y}_{obs})$ is gaussian.



(a) Full Bayes (b) MAP approximation (c) FMP approximation

Unimodal Gaussian Posterior

Let us assume that the joint posterior p($\boldsymbol{\theta}, \psi | \mathbf{y}_{obs}$) is gaussian.

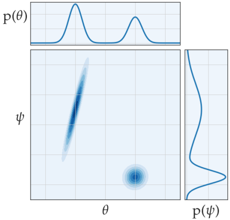


(a) Full Bayes (b) MAP approximation (c) FMP approximation

- The FMP method is exact on this case. With traditional MAP the parameter posterior shows reduced variance ("false certitude").

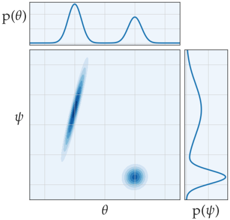
Mixture of Gaussians

Suppose that p($\boldsymbol{\theta}, \psi | \mathbf{y}_{obs}$) is a mixture of gaussians, well-separated.

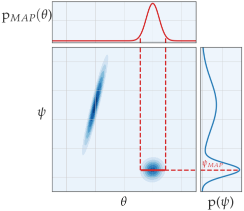


(a) Full Bayes

Suppose that p($\theta, \psi | \mathbf{y}_{obs}$) is a mixture of gaussians, well-separated.



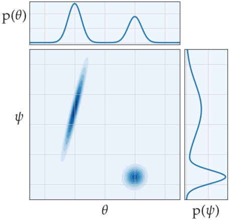
(a) Full Bayes



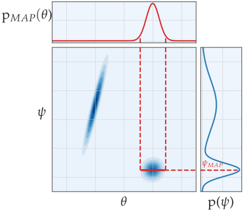
(b) MAP approximation

Mixture of Gaussians

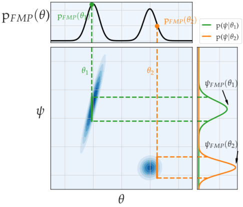
Suppose that $p(\boldsymbol{\theta}, \psi | \mathbf{y}_{obs})$ is a mixture of gaussians, well-separated.



(a) Full Bayes



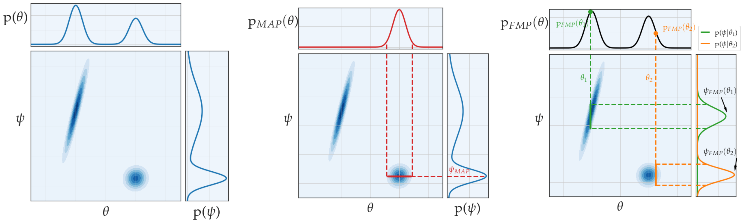
(b) MAP approximation



(c) FMP approximation

Mixture of Gaussians

Suppose that p($\boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{y}_{obs}$) is a mixture of gaussians, well-separated.



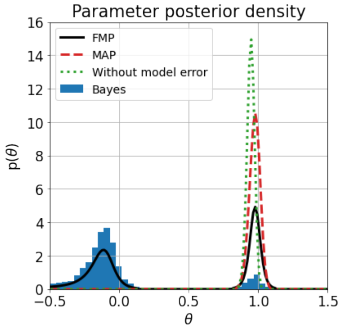
(a) Full Bayes (b) MAP approximation (c) FMP approximation

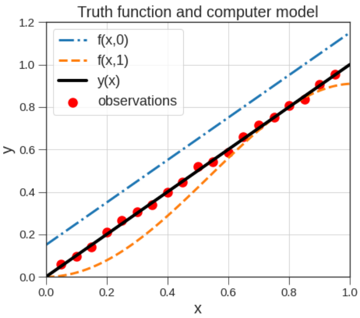
- Traditional MAP might not see probability mass. The estimated variance is still reduced.
- FMP approximation finds all maxima with correct variance. Peak values might be incorrect due to volume effects.

Contents

Numerical example

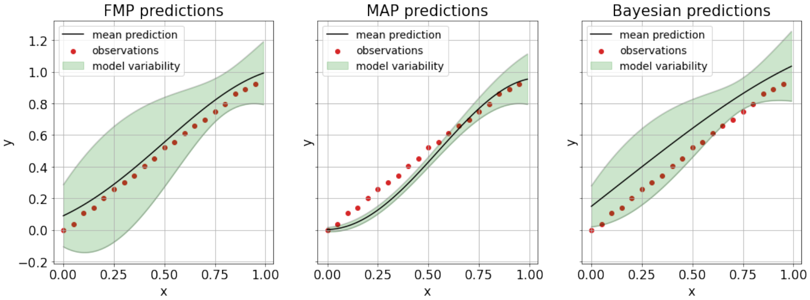
1D example. $y_{truth}(x) = x$ and $f(x, \theta) = x \sin(2\theta x) + (x + 0.15)(1 - \theta)$.
$x \in [0, 1]$ and $\theta \in [-0.5, 1.5]$.





We use 20 observations from reality, with small noise $\sigma^2_{exp} = 0.01$.
Statistical assumptions : $p(\sigma^2, \sigma^2_{exp}) \propto 1$, $l_{cor} \sim IG(5, 0.4)$, $\mu = 0$.
$p(\theta) \propto 1$, squared exponential kernel.

Predictions with the calibrated model



Confidence intervals from FMP and Bayesian predictions are large enough to contain observations.
Map predictions provide an incorrectly thin confidence interval, ” false certitude”.

Contents

Conclusion

- We solve the joint estimation problem that is overlooked in litterature because :
  - it is expensive to solve
  - it doesn't match with the traditional definition of model error.
- The FMP estimation reduces the cost and is still accurate with gaussian posteriors.
- The traditional MAP estimation might miss posterior probability mass.

Perspectives :

- Application to more complex cases,
- Study of smart sampling techniques to increase the dimension of the problem.

References I

Marc C. KENNEDY et Anthony O'HAGAN. "Bayesian Calibration of Computer Models". In : *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 63.3 (2001), p. 425-464. ISSN : 1467-9868. DOI : 10.1111/1467-9868.00294. URL : https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00294 (visité le 14/02/2020).

Matthew PLUMLEE. *Bayesian Calibration of Inexact Computer Models : Journal of the American Statistical Association : Vol 112, No 519*. 2017. URL : https://www.tandfonline.com/doi/full/10.1080/01621459.2016.1211016 (visité le 14/02/2020).

References II

Heng XIAO et Paola CINNELLA. "Quantification of Model Uncertainty in RANS Simulations : A Review". In : *Progress in Aerospace Sciences* 108 (1er juil. 2019), p. 1-31. ISSN : 0376-0421. DOI : 10.1016/j.paerosci.2018.10.001. URL : http://www.sciencedirect.com/science/article/pii/S0376042118300952 (visité le 14/02/2020).

Thanks for your attention!